



TriDoshax: AI Based Panchakarma Treatment and Diet Recommendation System

Sujal Wadkar¹, Poonam Dhamal², Shreyans Magdum³, Vishwajeet Ghadge⁴

Student, Dept. of Information Technology, G H Rasoni College of Engineering and Management, Pune, India^{1,3,4}

Assistant Professor, Dept. of Information Technology, G H Rasoni College of Engineering and Management, Pune, India²

Abstract: Ayurvedic medicine offers a prevention-oriented, constitution-based approach to personalized healthcare through the Tridosha framework—Vata, Pitta, and Kapha—collectively describing an individual’s Prakriti. Despite its clinical relevance, widespread adoption is hindered by the scarcity of trained practitioners, the inherent subjectivity of manual assessments, and the absence of scalable digital tools. This paper presents TriDoshax, a web-based Clinical Decision Support System (CDSS) that implements Ayurvedic Prakriti classification through a hybrid architecture combining supervised machine learning with rule-based Ayurvedic scoring. A structured questionnaire capturing 26 physiological, psychological, and lifestyle attributes is used to collect user data, which is subsequently processed through an ensemble Random Forest classifier trained on approximately 5,000 records. The Random Forest model achieves an accuracy of 0.826, precision of 0.851, recall of 0.826, and F1-score of 0.812, outperforming Decision Tree, Support Vector Machine, and Logistic Regression baselines. The predicted Dosha type drives a personalized recommendation engine that suggests Panchakarma therapies, dietary plans, and lifestyle guidelines consistent with classical Ayurvedic principles. Additionally, a Retrieval-Augmented Generation (RAG) pipeline powers a conversational AI assistant capable of responding to Ayurveda-specific queries. TriDoshax bridges traditional Ayurvedic principles with modern AI-assisted healthcare delivery, providing an accessible platform for preventive health management.

Keywords: Ayurveda; Tridosha; Prakriti Classification; Panchakarma; Random Forest; Clinical Decision Support System; Retrieval-Augmented Generation; Personalized Healthcare.

I. INTRODUCTION

Ayurveda, one of the world’s oldest codified medical systems, adopts a fundamentally personalized approach to health through the Tridosha theory. The three biological humors—Vata (governing kinetic functions), Pitta (governing metabolic processes), and Kapha (governing structural stability)—interact to define an individual’s Prakriti, or constitutional type [1]. Prakriti influences physiological predispositions, disease susceptibility, and therapeutic responses, making its accurate identification essential for effective Ayurvedic care.

Despite its theoretical richness, the practical adoption of Ayurveda faces systemic barriers. Prakriti assessment traditionally requires expert consultation involving the nuanced evaluation of physical, physiological, and behavioral traits—a process that is both time-intensive and difficult to standardize at scale. The limited availability of qualified Ayurvedic practitioners leads many individuals to rely on informal online resources, which may result in inaccurate self-assessments and suboptimal health decisions. A critical review of 64 existing Prakriti assessment instruments has revealed significant inconsistencies across tools and the absence of universally accepted evaluation standards [12].

Concurrently, advances in Artificial Intelligence (AI) and Machine Learning (ML) have created a compelling opportunity to formalize and scale Ayurvedic assessment. Machine learning models trained on structured questionnaire data can learn classification boundaries between Dosha types, supporting a more standardized and reproducible assessment process [7]. The growing availability of Prakriti-annotated datasets further enables rigorous model training and validation [11].

This paper introduces TriDoshax, an AI-enabled CDSS that integrates a Random Forest-based Prakriti classifier with a rule-based Dosha scoring engine and a RAG-powered conversational assistant. The system delivers an end-to-end pipeline—from structured user input and automated Dosha prediction to personalized Panchakarma therapy, dietary, and lifestyle recommendations. The key contributions of this work are: (i) a hybrid Dosha prediction framework combining ensemble ML with rule-based Ayurvedic scoring; (ii) a 26-attribute structured questionnaire encoding physiological, psychological, and lifestyle markers; (iii) a personalized recommendation engine mapping Prakriti types to Panchakarma



therapies, dietary protocols, and lifestyle guidelines; (iv) a RAG-based conversational assistant for Ayurveda-specific queries; and (v) a deployed web-based CDSS built on React, FastAPI, and PostgreSQL.

II. PROBLEM STATEMENT

Ayurveda is a prevention-oriented healthcare system that tailors interventions to an individual's biological constitution rather than applying generalized treatment protocols. However, realizing this potential at scale requires reliable Prakriti identification—a task presently constrained by practitioner scarcity and methodological inconsistency.

In the absence of accessible expert consultation, individuals increasingly turn to informal digital sources for self-assessment. These sources lack domain grounding, vary widely in the traits they assess, and offer no mechanism for personalized follow-up recommendations. Such misclassification may reduce the relevance of recommendations and potentially reinforce incorrect health behaviors. A standardized, AI-assisted digital platform capable of providing structured and consistent Prakriti assessment is therefore needed to expand the reach of Ayurvedic preventive healthcare.

III. LITERATURE REVIEW

The convergence of Ayurveda with computational methods has gained increasing scholarly attention. Early foundational work established machine learning architectures for Ayurvedic diagnosis [15], while subsequent studies demonstrated the feasibility of classification and drug-recommendation systems grounded in Ayurvedic principles [9][5]. AI-driven chatbot systems such as PrakritiBot [10] have shown promise for conversational Prakriti analysis, and the Prakriti200 dataset [11] has provided a structured benchmark for questionnaire-based model training. Research incorporating genetic markers [8] further strengthens the scientific foundation for Dosha-based personalization.

Despite this progress, several limitations remain. First, most existing systems employ either machine learning or rule-based logic in isolation; hybrid approaches that preserve Ayurvedic interpretability while incorporating predictive modeling remain underexplored [2]. Second, most deployed tools lack end-to-end pipelines connecting Dosha prediction to actionable recommendations including Panchakarma therapy. Third, the application of Retrieval-Augmented Generation for Ayurveda-specific conversational support is largely uninvestigated. TriDoshaX addresses these gaps by combining ensemble classification, rule-based scoring, personalized recommendation generation, and RAG-based dialogue into a unified CDSS.

IV. PROPOSED SYSTEM

A. Dataset and Preprocessing

The study utilizes an Ayurvedic Prakriti dataset comprising approximately 5,000 records, each representing a unique individual's questionnaire response profile. The dataset encodes 26 attributes spanning physiological descriptors (body frame, skin texture, digestion pattern, sleep quality), psychological traits (emotional tendency, stress response, memory), and lifestyle characteristics (energy patterns, appetite regularity, activity level). Each attribute is coded with categorical options (A, B, C) corresponding to Vata, Pitta, and Kapha trait expressions, respectively. Ground-truth labels indicate the dominant Dosha for each record.

Preprocessing comprised four stages. First, categorical responses were converted to ordinal numeric values via label encoding. Second, missing or inconsistent entries were identified and imputed using modal substitution to preserve dataset integrity. Third, feature normalization was applied to continuous-valued attributes to ensure scale invariance. Finally, the dataset was partitioned into training (80%) and evaluation (20%) subsets using stratified sampling to maintain class distribution across splits.

B. Machine Learning Models

Four supervised classifiers were evaluated: Random Forest, Decision Tree, Support Vector Machine (SVM), and Logistic Regression. These were selected to represent a spectrum of model complexity—from linear discriminants to non-parametric ensembles—enabling systematic comparison of their suitability for multi-class Prakriti classification.

1) Random Forest for Dosha Classification:

Random Forest is an ensemble method that constructs a collection of decorrelated decision trees through bootstrap aggregation (bagging) and random feature subspace selection. For a dataset $D = \{(x_i, y_i)\}_{i=1}^k$ with feature vectors $x_i \in \mathbb{R}^{26}$ and class labels $y_i \in \{\text{Vata, Pitta, Kapha}\}$, each tree T_k is trained on a bootstrap sample $D_k \subset D$. At each node, a random subset of $m = \sqrt{p}$ features (where $p = 26$) is evaluated for optimal splitting, reducing inter-tree correlation and variance.



The final prediction is aggregated via majority voting:

$$\hat{y} = \arg \max_c \sum_{k=1}^K \mathbb{1}[T_k(x) = c]$$

where $K = 100$ denotes the number of trees. This ensemble mechanism reduces overfitting and improves classification stability compared to individual decision trees [5]. The model was configured with 100 estimators, a fixed random seed for reproducibility, and a constrained maximum depth to prevent overfitting. Training was performed using Scikit-learn in a Google Colab environment.

C. Hybrid Scoring Architecture

TriDoshax employs a two-stage prediction strategy. The Random Forest model determines the primary Dosha type along with associated class probabilities. A complementary rule-based scoring module, informed by classical Ayurvedic trait-to-Dosha mappings, computes secondary Dosha scores and confidence levels by aggregating weighted attribute responses. This hybrid design ensures that predictions remain consistent with established Ayurvedic principles, improving interpretability and reducing the risk of logically inconsistent recommendations. The integration of data-driven ML with domain-grounded rules reflects established best practices in hybrid CDSS design.

D. RAG-Based Conversational Assistant

The system incorporates a Retrieval-Augmented Generation (RAG) pipeline to support open-ended Ayurvedic queries. A curated corpus of Ayurvedic texts and domain guidelines is indexed using a dense embedding model, with semantic retrieval performed against a vector database. Retrieved context passages are injected into the prompt of a large language model, enabling it to generate more grounded and contextually relevant responses. This approach reduces the hallucination risks inherent in purely generative models when applied to domain-specific queries, and produces more informative responses than static FAQ-based approaches.

V. SYSTEM ARCHITECTURE

The TriDoshax system is organized into six functional layers, as illustrated in Fig. 1.

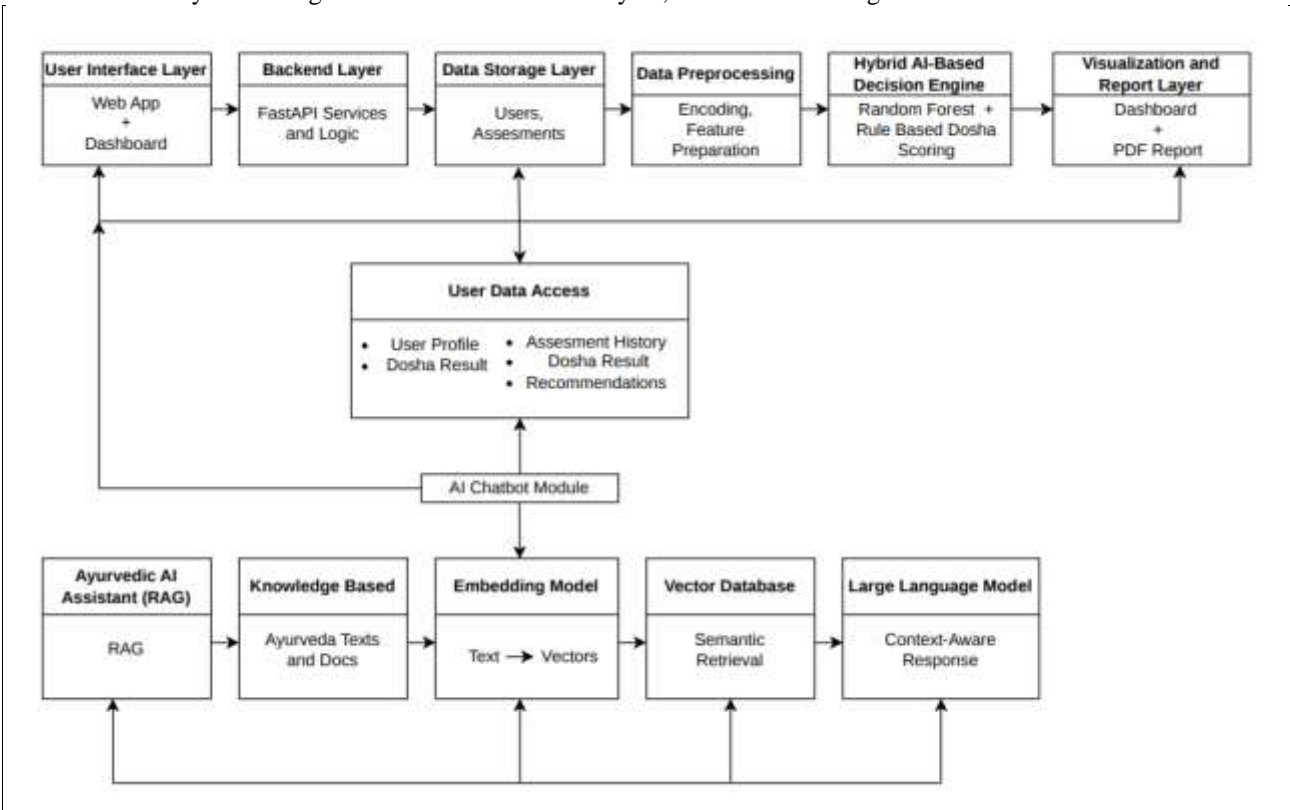


Fig. 1 System Architecture of the TriDoshax Platform

User Interface Layer: A React-based web application with Tailwind CSS provides a responsive, accessible interface including the Prakriti questionnaire, wellness dashboard, and chatbot interaction panel.



Backend Layer: A FastAPI service layer manages request routing, business logic execution, and ML model inference. RESTful endpoints expose assessment, prediction, and recommendation functions.

Data Storage Layer: PostgreSQL stores user profiles, assessment responses, Dosha results, historical records, and recommendation artifacts in a structured relational schema.

Data Preprocessing and Decision Engine: Incoming questionnaire responses undergo label encoding and normalization before being passed to the Random Forest inference pipeline. The hybrid decision engine combines ML-predicted probabilities with rule-based Dosha scores to produce the final Prakriti classification and confidence metrics.

Visualization and Report Layer: The wellness dashboard presents Dosha distribution charts, disease susceptibility indicators, and structured recommendation panels. Users may download a generated PDF health report for offline use.

Ayurvedic AI Assistant (RAG): Ayurvedic knowledge sources are encoded into vector embeddings, stored in a retrieval database, and queried at runtime to provide the language model with relevant context for each user query.

VI. RESULTS AND DISCUSSION

A. Experimental Setup

All experiments were conducted using Python 3.10 with Scikit-learn for model training and evaluation. The Random Forest model was trained in Google Colab and deployed via FastAPI. The web application was tested on a system with an Intel i5/i7 processor and 8–16 GB RAM, with PostgreSQL as the data persistence layer.

B. Classification Performance

Table I reports the performance of all evaluated models on the held-out test set (20% of 5,000 records).

TABLE I CLASSIFICATION PERFORMANCE OF MACHINE LEARNING MODELS

Algorithm	Accuracy	Precision	Recall	F1-Score
Random Forest	0.826	0.851	0.826	0.812
Decision Tree	0.771	0.774	0.771	0.772
SVM	0.701	0.716	0.701	0.687
Logistic Regression	0.638	0.648	0.638	0.631

Random Forest achieves the best performance across all metrics. Its 5.5 percentage-point accuracy lead over the Decision Tree is attributable to ensemble aggregation, where individual trees collectively reduce variance through majority voting. SVM (0.701) and Logistic Regression (0.638) underperform significantly, confirming that the Dosha feature space is inherently non-linear and cannot be adequately separated by hyperplane-based models. The low F1-score of Logistic Regression (0.631) further reflects poor generalization to minority Dosha classes.

The macro-averaged precision of 0.851 reflects consistent performance across all three Dosha classes, while the F1-score of 0.812 indicates balanced precision-recall behavior—a desirable property in multi-class health-related classification tasks. Fig. 2 visualizes these comparative results across all four models.

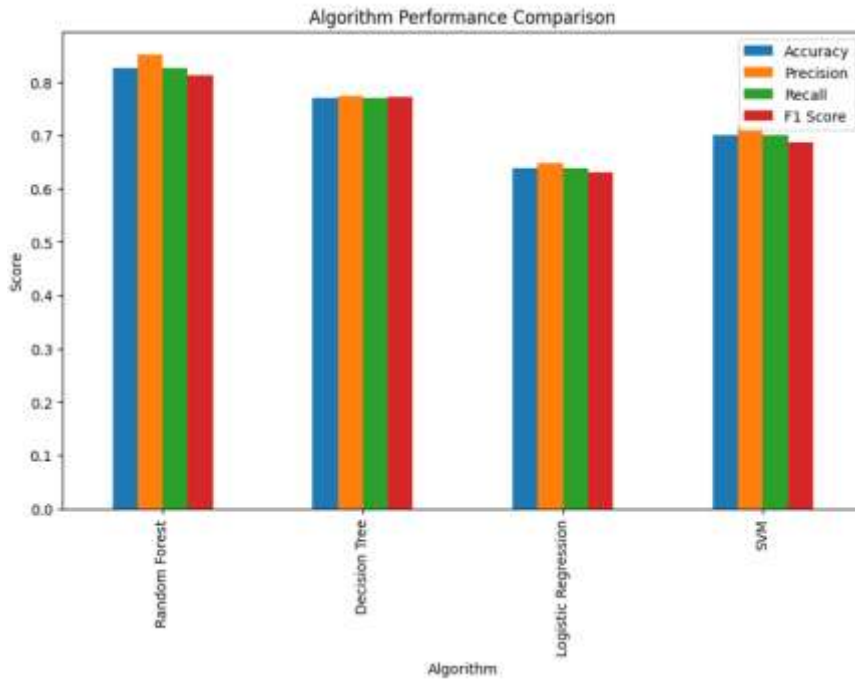


Fig. 2 Comparative Classification Performance of Machine Learning Models

C. Confusion Matrix Analysis

The confusion matrix in Fig. 3 provides per-class insight into classifier behavior. The Kapha class achieves strong classification, with 414 of 441 instances correctly identified, producing only minor misclassification into Pitta (25) and Vata (2). Pitta performance is similarly strong, with 333 correct predictions and 34 misclassified as Kapha. These results indicate that the physiological and behavioral attributes associated with Kapha and Pitta are sufficiently discriminative within the questionnaire feature space.

The Vata class presents a greater classification challenge: only 79 instances are correctly identified, while 107 are misclassified as Kapha and 6 as Pitta. This pattern is consistent with the Ayurvedic understanding that Vata and Kapha doshas share several overlapping trait expressions—which reduce inter-class separability in the feature domain. Targeted data augmentation for Vata-dominant profiles and the incorporation of pulse diagnostics or phenotypic biomarkers represent promising directions for improving Vata classification in future iterations.

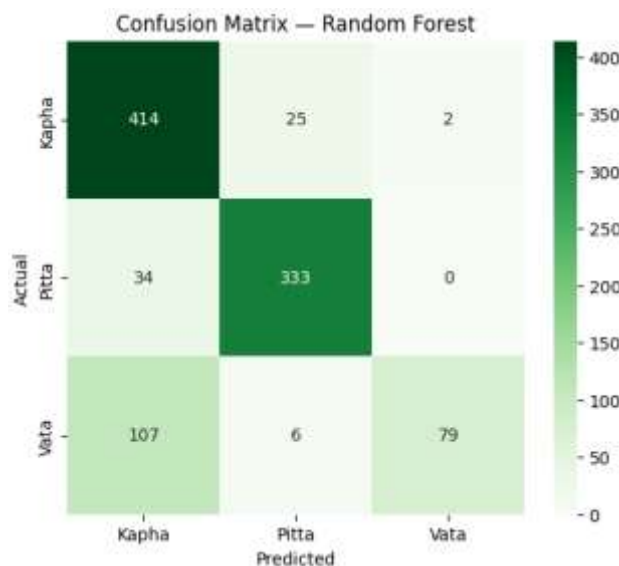


Fig. 3 Confusion Matrix of the Random Forest Classifier



D. System Output and Dashboard

The wellness dashboard (Fig. 4) integrates predictive outputs with structured recommendations. The primary panel displays the predicted Dosha (e.g., Pitta) with a confidence score (52%) and secondary Dosha identification. A donut chart visualizes the proportional Dosha distribution. Disease susceptibility indicators—color-coded across digestive health, inflammation risk, respiratory function, and stress tolerance—provide contextual guidance. Recommendation panels deliver personalized dietary guidance, lifestyle suggestions, and Panchakarma therapy recommendations. Users may download a structured PDF health report or retake the assessment to track changes over time.

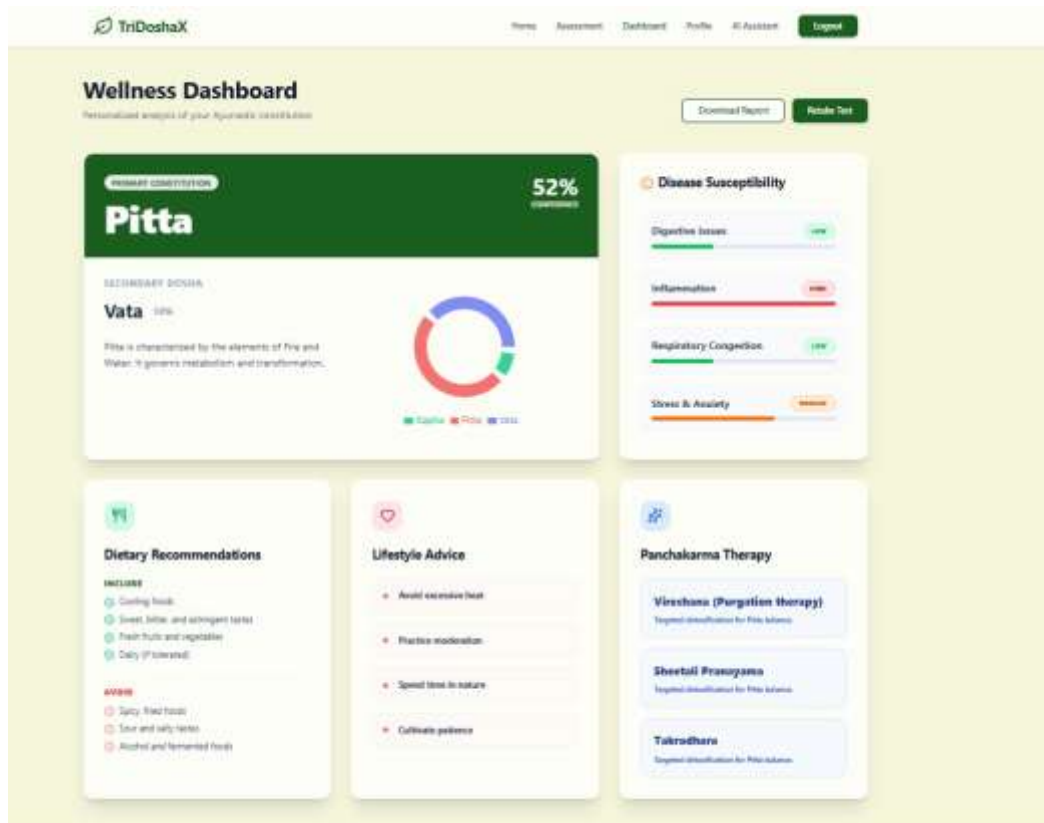


Fig. 4 TriDoshax Wellness Dashboard

E. Discussion

The results confirm that ensemble ML applied to structured Ayurvedic questionnaire data achieves meaningful Prakriti classification performance. Random Forest's superiority over linear and single-tree models reflects the non-linear, multi-dimensional nature of the Dosha feature space.

The hybrid architecture—combining ML prediction with rule-based scoring—is particularly valuable in this context. Pure ML outputs may occasionally conflict with established Ayurvedic principles; the rule-based layer anchors recommendations to validated classical guidelines, improving interpretability and user trust. This approach aligns with established best practices in hybrid CDSS design.

The RAG pipeline addresses a critical usability need: users will frequently pose questions outside the scope of the questionnaire. Grounding language model responses in curated Ayurvedic knowledge reduces hallucination and enables more informative, contextually appropriate responses compared to static FAQ-based approaches.

VII. CONCLUSION

This paper presented TriDoshax, a hybrid AI-based Clinical Decision Support System for Ayurvedic Prakriti classification and personalized health recommendation. By integrating a Random Forest ensemble classifier with rule-based Dosha scoring, a structured recommendation engine, and a RAG-powered conversational assistant, the system delivers structured, interpretable Ayurvedic health guidance through an accessible web interface.

The Random Forest classifier achieves 82.6% accuracy and 85.1% precision on a 5,000-record questionnaire dataset, outperforming all evaluated baselines. Confusion matrix analysis reveals strong classification performance for Kapha



and Pitta doshas, with Vata-Kapha confusion identified as the primary remaining challenge, motivating targeted dataset expansion and feature enrichment in future iterations.

Key limitations include the reliance on self-reported questionnaire responses (which may introduce recall bias), the absence of validation against practitioner-assigned Prakriti labels, and potential underrepresentation of Vata-dominant profiles in the training data. Future work will pursue practitioner-validated dataset curation, multi-modal input incorporating pulse diagnostics and phenotypic biomarkers, longitudinal Prakriti tracking to capture seasonal and lifestyle-driven shifts, and evaluation of recommendation outcomes in real-world settings.

TriDoshax demonstrates that AI integration can make Ayurvedic preventive healthcare more accessible and standardized, contributing to the growing field of AI-assisted personalized health systems.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Department of Information Technology, G H Raison College of Engineering and Management, Pune, for providing research infrastructure and support throughout this work.

REFERENCES

- [1]. S. K. Tiwari, R. R. Mahto, H. Sharma, P. Khandelwal, and S. Chauhan, "Ayurveda and Artificial Intelligence: A Review of Applications in Diagnosis, Prakriti Analysis, and Personalized Therapeutics," *J. Health Synapse*, vol. 1, no. 1, pp. 24–35, 2026.
- [2]. M. Ranade, "Artificial Intelligence in Ayurveda: Current Concepts and Prospects," *J. Indian Syst. Med.*, vol. 12, no. 1, pp. 53–59, 2024.
- [3]. V. Biradar, N. Patel, and J. Rathod, "Automation and Robotics in Panchakarma Therapy: A Conceptual Proposed Review on Modernizing Traditional Ayurvedic Practices," *World J. Pharm. Pharm. Sci.*, vol. 14, no. 11, pp. 2428–2437, 2025.
- [4]. H. M. Manjula and S. P. AnandaRaj, "Ayurvedic Diagnosis using Machine Learning Techniques to Examine Diseases by Extracting Data Stored in AyurDataMart," in *Proc. 3rd Int. Conf. Advances in Computing, Communication Control and Networking (ICACCCN)*, IEEE, 2021, pp. 239–243.
- [5]. S. Kumar, H. Kumar, N. Kumar, and M. Das, "Ayurvedic-Based Disease Prediction and Drug Recommendation System," *Srusti Manag. Rev.*, Special Issue, pp. 27–37, 2024.
- [6]. V. Rao et al., "Establishing Key Components of a Combined Ayurvedic Diet and Yoga Therapy Program for Weight Management in Women with Polycystic Ovary Syndrome: A Delphi Study," *BMC Complement. Med. Ther.*, vol. 25, p. 402, 2025.
- [7]. S. K. Verma, M. Pandey, A. Sharma, and D. Singh, "Exploring Ayurveda: Principles and Their Application in Modern Medicine," *Bull. Natl. Res. Cent.*, vol. 48, p. 77, 2024.
- [8]. M. Murugan et al., "Genetic Variants of Leptin Receptor Gene (rs1137101) and Obesity Risk in Prakriti Individuals and Its Pathogenicity Prediction Using In Silico Approaches," *Egypt. J. Med. Hum. Genet.*, vol. 26, p. 92, 2025.
- [9]. V. S. Jadhav, A. D. Wakale, and S. R. Mane, "Integration of Machine Learning in Ayurveda: An Indian Traditional Health Science," *Int. Res. J. Sci. Eng.*, Special Issue A14, pp. 57–62, 2024.
- [10]. K. Hemalatha et al., "Prakriti Bot – Intelligent Prakriti Analysis Chatbot for Personalized Ayurvedic Suggestions," *IRJIET*, vol. 9, pp. 337–342, 2025.
- [11]. A. K. Singh and J. Singh, "Prakriti200: A Questionnaire-Based Dataset of 200 Ayurvedic Prakriti Assessments," *IEEE Data Descriptor*, DOI: 10.21227/k748-f379, 2024.
- [12]. A. Venkatesh et al., "Prakriti (Constitutional Typology) in Ayurveda: A Critical Review of Prakriti Assessment Tools and Their Scientific Validity," *Front. Med.*, vol. 12, p. 1656249, 2025.
- [13]. R. Tiwari et al., "Scopes and Uses of Artificial Intelligence in the Field of Ayurveda: A Review Study," *Afr. J. Biomed. Res.*, vol. 27, no. 3s, pp. 3597–3603, 2024.
- [14]. S. Khan, A. Saify, S. Gosaliya, D. Jain, and J. Zalte, "MediMatch: AI-Driven Drug Recommendation System," *Shah and Anchor Kutchhi Engineering College Research Publication*, 2023.
- [15]. "Review Articles on Artificial Intelligence Applications in Ayurveda," *J. Ayurveda Integr. Med. Sci. (JAIMS)*, vol. 9, no. 10, ISSN 2456-3110, 2024.